Reproductive success and survival during the breeding season in relation to individual behaviour in the great tit, *Parus major*

Camilla Thorsteinsen

CEES

Centre for Ecological and Evolutionary Synthesis

Master of Science Thesis
Centre for Ecological and Evolutionary Synthesis
Department of Biosciences

UNIVERSITY OF OSLO

01.06.2015

2015

Reproductive success and survival during the breeding season in relation to individual behaviour in the great tit, *Parus major*

Camilla Thorsteinsen

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo

Abstract

Repeatable behavioural variation within individuals that is consistent over time and across contexts is often defined as non-human animal personality. Individuals may be classified as having a shy or bold personality. Shy individuals typically react to unfamiliar situations with reactivity or avoidance and bold individuals with proactivity or aggression. When studying a population, accounting for variation in boldness may for instance prevent sample biasing or help explain observed trade-offs or non-optimal behaviour. Relationships between fitness and personality may exist in many populations but previous studies have studied only one or a few behavioural traits and the results differ substantially. This study explored possible relationships between reproductive success or survival during the breeding season and behavioural traits recorded in the wild in a population of great tits, *Parus major,* and is the first to obtain results with as many recorded behavioural traits for the same bird.

It is unclear what is maintaining variation in personalities in a population but one explanation may be that the personalities have equal fitness over time because a trade-off between survival and reproductive success exists. Another explanation may be that all individuals exhibit the best possible behavioural type given their condition but one phenotype will be superior and have both higher reproductive success and survival. To determine which explanation was most likely, measured behavioural traits were classified as either bold or shy using previous studies results or logic reasoning. Earlier studies indicated that bold birds had higher reproductive success than shy birds, which was supported by the present study. However, no relationship between survival during the breeding season and behaviour was obvious from the previous or current study, which made the determination of whether one explanation was more likely than the other problematic. Further studies should use data from several years as the effect of personality may vary between years or seasons, test repeatability for more traits to determine whether they represent personality, and develop and use more standardised traits. In addition to age, sex and breeding time more confounding variables is probably also important to account for, for instance temperature, weather conditions, date and time of day.

Acknowledgements

This master thesis was carried out at CEES (Centre for Ecological and Evolutionary Synthesis) at the Department of Biosciences, University of Oslo. Prof. Tore Slagsvold and Prof. Helene M. Lampe have supervised this project.

- I wish to thank my main supervisor Tore Slagsvold for giving me the opportunity to work with him and others during my two years at CEES. It has been a great time, and I am grateful for the continuous and invaluable guidance I have received.
- My good friend and fellow master student Tina Øiseth Stræte deserves many hugs for helping me with the fieldwork and work that followed, and for fun times spotting snakes, moose and birds at Dæli.
- I also wish to thank my co-supervisor Helene M. Lampe for good feedback on the draft.
- I thank Anders Herland for assistance in the field and for transferring raw data from our rained on, muddy fieldwork papers into Excel files.
- Finally, I thank my family, friends and boyfriend for being supportive and wonderful. I also want to thank you for reminding me that there is a world outside of Blindern (who knew!). I love you all.

Table of contents

1 Introduction

Inter-individual differences in behaviour that are consistent over time and across contexts, with contexts defined as functional behavioural categories (e.g., feeding or mating), may be defined as for instance personality, coping styles or behavioural syndromes (Sih *et al.* 2004a; Smith & Blumstein 2007). From here on, the term "personality" is used. In non-human animal (henceforth referred to as "animal") populations repeatable behavioural variation within individuals is often taken as evidence of animal personality (Boulton *et al*. 2014). Personality is studied in a wide range of animals, both in model and non-model species (Sih *et al.* 2004a), and includes different apes and monkeys, rodents, reptiles, fish and bird species, illustrated in for instance Gosling's (2001) comprehensive review of the diverse animal personality literature. I have studied possible relationships in the field between a number of personality, or behavioural, traits and fitness measures in a small passerine bird, the great tit, *Parus major.*

1.1 Why study personality?

Animal personality may be studied for a number of reasons. First, it may help researchers approach or reach answers of difficult questions in human psychology studies, for instance concerning the context and evolution of human personalities (Groothuis & Carere 2005), as it might be easier for instance to investigate the biological, genetic or environmental influence on personality in animals than in humans (Gosling & Vazire 2002). Second, Sih *et al.* (2004a) suggested that behavioural traits should not be studied alone but together as a package because personality may greatly influence fitness (see also Smith & Blumstein (2007)), speciation rates, species distribution or invasion. In behavioural ecology, it usually has been assumed that all individuals in a population may have the capability to display an apparently optimal behaviour in all situations. However, the concept of personality implies that there may be a limit to behavioural plasticity, simply by suggesting that an individual's behaviour may be beneficial in some contexts but unsuitable in other contexts. This may help explain trade-offs or non-optimal behaviour observed in populations (Sih *et al.* 2004a). Third, individual differences in behaviour may lead to some individuals in a population being easier to detect and trap than others, producing a sampling bias (Wilson *et al*. 1994). Because population sizes often is decided using capture-recapture methods based on the assumption that all individuals have equal probability of being captured and recaptured, this may lead to

erroneous estimations of population sizes (Andersen 2012). Fourth, as reactions to stress or artificial housing may vary between individuals or species, a more comprehensive overview over individual differences in behaviour may benefit animal welfare. Lastly, individual variation in behaviour may include different responses towards experimental treatments. Hence, the interpretation or design of studies and experiments may benefit from animal personality studies (Groothuis & Carere 2005).

In other words, many benefits may be drawn from this field of research, but to fully appreciate them depends on the tenability of the animal personality concept. Gosling & Vazire (2002) reviewed the literature and found strong evidence for animal personality not reflecting researchers' imagination. Tendencies revealed strong agreements in the assessment of personality, and the structure of personality ratings were concluded, to some degree, to predict real quantities of studied individuals. Overall, it was concluded that animal personality really exists.

1.2 The shy-bold continuum

Some commonly studied behavioural traits are for instance exploration, how an individual explores novel objects, environments or situations, and aggression (see for instance Gosling 2001). Another example is hissing, thought to be an anti-predator behaviour in birds (Krams *et al.* 2013b), and breathing rate, which may be a measure of stress or anxiety (Class *et al.* 2014). In the recent years, scientists have studied behaviour in the form of shy- or boldness in a variety of species, and today the so-called shy-bold continuum is indeed a well-documented and important axis of behavioural variation (Wilson *et al*. 1994). On this axis, extremely shy individuals will be located at one far end of the axis, identified by typically reacting to unfamiliar situations with avoidance and becoming alert and reactive. Extremely bold individuals, which will act aggressive, actively exploratory or interested in the same situations, will be found at the other far end of the axis (Wilson *et al*. 1993). Hence, the bold individuals may be easier to detect and trap than the shy individuals, which almost by definition may be prone to becoming an invisible, hidden segment of the study population (Wilson *et al*. 1994).

1.3 Development of individual behaviour

The psychology literature presents two alternative views on the development of individual behaviour, the context general and the context specific view (Colléter & Brown 2011; Sih *et al.* 2004a). The context general view proposes that populations in different environments should produce similar correlations between traits because traits are generally constrained and often correlated (Colléter & Brown 2011). Consequently, this view is compatible with the animal personality framework, which implies limited phenotypic plasticity by classifying individuals as more or less aggressive, bold, exploratory, and so on (Sih *et al.* 2004a). A study by Malmkvist & Hansen (2002) on farm mink, *Mustela vison*, demonstrated this view. Minks selected for fearfulness against humans over 10 generations generalized their fear and were more fearful towards both humans, an object and other minks than minks selected for confidence against humans for 10 generations. The alternative context specific view proposes that independent selection on individual traits may be present, allowing for greater uncoupling of traits and larger degrees of phenotypic plasticity. Coupling of traits will occur only under certain contexts (Colléter & Brown 2011). Context specificity has been found in for instance pumpkinseed sunfish, *Lepomis gibbosus.* Individual fish that consistently presented bold responses towards a threatening object did not necessarily do the same towards a novel food object, suggesting that individual differences in boldness and shyness may be context dependent. (Coleman & Wilson 1998)

1.4 Maintaining variation within a population

Little is known about the genetic correlations between behavioural traits on the shy-bold axis (Sih *et al.* 2004b) and how they are maintained within populations. One adaptive explanation is that it reflects life-history variation (Cole & Quinn 2014), i.e., variation in traits directly associated with reproduction and survival (Stearns 1992). To maintain variation in behavioural traits in a population (i.e., for natural selection to maintain the shy-bold axis) one genotypic and one phenotypic condition must be met. The genotypic condition considers the heritability of variation in traits, which will be a response to selection (Stearns 1992). Heritability is demonstrated for several behavioural traits, including breathing rate, aggression (Class *et al*. 2014) and exploration (Dingemanse *et al*. 2002). The phenotypic condition contemplates that individuals must vary in fitness, e.g., in the total lifetime reproductive output or survival (Stearns 1992).

Variation in traits may be maintained because the fitness of personalities depends on the context they are expressed in. For instance, being bold or aggressive may be beneficial in contexts with high competition for territories or mates, as it may increase reproductive success, but disadvantageous in contexts with high predation rates, as it may decrease survival rate (Smith & Blumstein 2007). This suggests that trade-offs in fitness consequences across contexts may sustain behavioural variation in populations (Sih *et al.* 2004b). Because the environment is variable the bold individuals may do well in some contexts and the shy in other contexts, and the personalities may represent different evolutionary stable strategies (ESS) (Sih *et al.* 2004b). Hence, the shy and bold personality may be maintained in a population because they have similar long-term fitness (Smith & Blumstein 2007). A meta-analysis on fitness consequences of animal personalities by Smith & Blumstein (2007) supported that bold individuals might have higher reproductive success, especially males, but lower survival rate than shy individuals due to such trade-offs.

An alternative hypothesis is the best-of-a-bad-job hypothesis where fitness is related to the ability to compete. Individuals with high competitive abilities (e.g., bold, large or strong individuals) will express their superior phenotype and, thus, have higher reproductive success and survival rate than individuals with low competitive abilities (e.g., shy, small or young individuals). Personality is considered plastic and all individuals will exhibit the best possible behavioural type given their own condition and ability. Hence, individuals that do not inhabit the superior phenotype must cope with a best-of-a-bad-job solution (Sih *et al.* 2004a).

1.5 Earlier studies on relationships between fitness and personality

As Gosling (2001) comprehensive review of animal personality studies described, the studies often concern only one or a few traits. Self-reporting by study subjects is possible only when studying humans, which might be one of the reasons explaining this bias. Many studies focused on broad dimensions of personality (e.g., sociability or emotionality) and attempted to relate the dimensions to specific behaviours or traits that were relatively narrow (e.g., attacks or social play). Behavioural traits were not always the primary focus of many personality studies. In fact, investigators whose primary interests were not in personality had conducted much of the research. By investigating the literature, I found that most of the studies tested only one or a few behavioural traits, mainly explorative behaviour and

aggression, and that few tested the relationships between the traits and survival during the breeding season (Table 1).

Table 1. A summary table over six personality studies where one or more personality traits were compared with various fitness traits.

Personality traits	Fitness measures	Correlation	Sex	Species	Reference
Breathing rate	Survival	-,o	F, M	Blue tit, *Cyanistes caeruleus*	Class *et al.* (2014)
Hissing	Survival	+	F	Great tit, *Parus major*	Krams *et al.* (2013)
	Clutch size	o	F		
Aggression	Survival	o	F, M	Blue tit	Class *et al.* (2014)
	Number of recruits	+	M, F*M		
Aggression	Parturition date	o	F	Red squirrel, *Tamiasciurus hudsonicus*	Boon *et al.* (2007)
	Litter size	o	F		
	Offspring survival	+/-	F		
Exploration	Survival	+/-, -/+	F, M	Great tit	Dingemanse *et al.* (2004)
	No. recruits	+/-, o	F, M		
Exploration	Survival	+	F	Great tit	Vrublevska *et al.* (2014)
	Nest success	+	F		
Exploration	Territory quality	+	M	Great tit	Both *et al.* (2005)
	Laying date	o	F, M		
	Clutch size	o	F, M		
	Prob. producing fledglings	-	F		
	Nest success	-	F		
	Fledgling size	-, o	F, M		
	No. fledglings	o	F, M		
	Fledgling condition	+/-	F*M		
Activity (exploration and locomotion)	Parturition date	o	F	Red squirrel	Boon *et al.* (2007)
	Litter size	o	F		
	Juvenile growth rate	+/-	F		

+: Positive correlation ($p < 0.05$)
-: Negative correlation ($p < 0.05$)
o: No correlation
/: Correlation differs between years or seasons
F: Female
M: Male
*: Interaction between the female and male behaviour
Note: For descriptions of variables, see main text.

In behavioural and personality studies, birds are commonly used as study species (see for instance Table 1) and the birds species that has probably received most attention is the great tit, a model species with a well-known behaviour and ecology (Groothuis & Carere 2005). Main reasons for choosing this species for personality studies are possibly that it is common in large areas of the world (Haftorn 1971), it breeds in nest boxes, can be bred in captivity and nestlings may be reared by hand (Groothuis & Carere 2005).

In studies where fitness and personality have been compared, the different fitness variables measured have been numerous, including for instance clutch size and juvenile growth rate. The results differ, for instance with exploration in great tits correlating negatively with nest success in Both *et al.'s* (2005) study, but correlating positively in Vrublevska *et al.'s* (2014) study. Reproductive measures have positive or negative relationships with personality traits measured, and no clear relationship between survival and personality has been found (Table 1). In some of the studies I investigated, evidence for the two hypotheses mentioned above was found. Vrublevska *et al.* (2014) found that survival during the breeding season and nest success correlated positively with the explorative behaviour of great tit females, supporting the best-of-a-bad-job hypothesis (Sih *et al.* 2004a) with the best phenotype being fast exploration. Dingemanse *et al.* (2004) found that the effect of exploration differed between years and sexes in the great tit, with fast explorers surviving better some years and slow explorers other years. A relationship between the number of recruiting young and female exploration that differed between years was also found. The researchers suggested that annual changes in selection pressures were a result of differences in food abundance from one year to the next. Hence, this study support the ESS hypothesis. In Both *et al.'s* (2005) study of great tits, nestlings of slow exploring females had higher probability of fledging and higher body mass than fast exploring females' nestlings. Exploration is highly associated with boldness, with slow explorers assumed shy and fast explorers assumed bold (Wilson *et al.* 1993). Hence, this study indicate that shy females have greater reproductive success than bold females. However, assortatively mated pairs produced offspring in better condition than dissimilar pairs. Hence, different selection pressures may have influenced the population, making shy- and boldness equally good phenotypes, as expected under the ESS hypothesis. In Boon *et al.'s* (2007) study of North American red squirrels, *Tamiasciurus hudsonicus*, the same hypothesis was supported. Offspring survival to the next breeding season was related to the female parent's aggression, and the effect varied between years. Likewise, juvenile growth rate was related to the female's activity (i.e., exploration and locomotion) and varied between

years. In Class *et al.'s* (2014) cross-fostering study of blue tits, *Cyanistes caeruleus,* a higher probability of recruiting was found for nestlings reared by aggressive males or assortatively mated pairs for aggression, indicating that bold birds have higher reproductive success than shy birds. Handling aggression was not associated with survival in either sex. In Krams *et al.'s* (2013) study of great tit females, survival during the breeding season was higher for females that hissed towards a predator than females not hissing. No relationship was found between clutch size and hissing. These results indicate that bold, hissing females have higher survival during the breeding season than shy females.

A previous study of the current population of great tits focused on the methodology of personality testing, with data from the 2010 and 2011 seasons (Andersen 2012). Andersen (2012) used the behavioural traits biting, screaming, alarm calling, breathing rate, tonic immobility and flight distance when holding a bird in the hand, as described in the method section below. Repeatability within observers was found to be relatively good and repeatability between observers was found to be relatively poor. Repeatability between the first and second independent measurement of traits was found for screaming and breathing rate. The strength of the relationships between behavioural traits was measured to determine whether the traits represented reliable methods for assessing personality. A few significant relationships were found and are discussed later, as well as relationships between various behavioural traits and fitness.

1.6 Present study

In the present study, I examine the possible relationships between individual personality and fitness (i.e., survival during the breeding season and reproduction) in a population of great tits. I will test predictions from the two different hypotheses mentioned above, and compare my results with earlier studies.

The great tit was considered a suitable species for studying behaviour in the study area for several reasons. First, it is one of the most common species used for personality studies (Groothuis & Carere 2005). Second, the great tit is abundant in our study area and nearly all of the breeding great tits use the nest boxes for breeding (Hansen *et al.* 2007). This simplifies tracking the reproduction and survival of individual marked birds that settle locally. Third, the great tit is a resident bird where most of the individuals stay in the area during the entire year (T. Slagsvold, unpublished data). This means that recruiting birds may be captured already in

autumn for personality testing. Fourth, I follow up a previous study of the same population in the same area (Andersen 2012). In the study by Dingemanse *et al.* (2004) mentioned above (Table 1) fitness consequences of behavioural traits in great tits varied across years. I will discuss differences and similarities between two breeding seasons (2011 and 2014) in great tits to see if the same pattern may be found in this study population. Fifth, few have studied personality and fitness in the wild (Smith & Blumstein 2007) and the existing results differ (see for instance Table 1). Additionally, few studies exist on relationships between survival during the breeding season and behavioural traits in great tits (however, see Krams *et al.* 2013 and Vrublevska *et al.* 2014).

To determine which behavioural traits may be classified as bold or shy, I have based my work on the findings by Andersen (2012) and on other studies, which will be discussed below.

1.6.1 Hypotheses and predictions

Mixed Evolutionary Stable Strategies (ESS) hypothesis: A trade-off between reproductive success and survival exist in the great tit population, and the shy and bold individuals represent different life-history strategies coexisting as mixed evolutionary stable strategies (ESS). Prediction: Shy birds will have lower reproductive success per breeding attempt but a higher survival rate than the bold birds. The total reproductive output will be equal for the two personalities

Best-of-a-bad-job hypothesis: Personality is plastic and all individuals express the best possible phenotype given their competitive ability. The shy personality is a best-of-a-bad job strategy, for instance caused by being reared in a poor environment, resulting in low phenotypic quality as an adult. Prediction: Shy individuals will have a lower reproductive success per breeding attempt and a lower survival rate than the bold birds.

2 Methods and materials

2.1 Study area and species

The study area was at Dæli in Bærum, near Oslo (60 00° N, 10 83° E). The Dæli area consists of a mixed deciduous-coniferous woodland that stretches across a 1.6km^2 area, containing approximately 500 nest boxes regularly distributed 40-50 m apart. Abundant birds using the nest boxes are great tits, blue tits, pied flycatchers, *Ficedula hypoleuca,* nuthatches, *Sitta europaea* and coal tits, *Periparus ater* (Slagsvold *et al.* 2013). Approximately 25 feeding sites are scattered throughout the study area (Andersen 2012). These sites are used for capturing and ringing birds in the autumn, and for behavioural and morphological measuring.

The study species was the great tit, a territorial passerine bird that has been studied at Dæli since 1995 (Hansen *et al*. 2007). Approximately 70-100 great tit pairs breed in the nest boxes every year. In the 2014 study season, no regular second clutches occurred after successful breeding but some individuals produced several clutches during the breeding season after a nesting failure. Only the birds' first clutches were used for the analyses and the dataset then contains information on clutches from 87 nests.

2.2 Capture and measurements

The age and gender of birds were decided after capturing based on the method described by Svensson (1975). Age was determined to be first year or older. All individuals got a uniquely numbered aluminium ring, while different combinations of colour rings illustrated to what category the birds belonged. In our study area, some great tit nestlings are raised by blue tit parents, and some blue tit nestlings are raised by great tit parents (see for instance Slagsvold *et al.* 2002). In the present study, data concerning cross-fostered birds were excluded. Nests where nestlings disappeared before fledging or died of other causes than starvation were assumed to be predated, and were excluded from the analyses. In 2014, this included five nests, which were considered insufficient for statistical analyses. The response to a predator model test (described below) was performed late in the nestling period (mean date 4th June ± 5 days). I was thus unable to compare the response to a predator model with survival because none of the tested birds died or disappeared after this test (Appendix table 3.1.a. and 3.2.a.).

The significance of personality on the males' disappearance was not tested, as only five males disappeared during the breeding season.

I focused on the females when running most of the analyses, as they may have the greatest impact on reproductive measures like laying date and clutch size. All personality tests were performed in the field.

2.3 Response to humans

2.3.1 In-hand personality test

During the autumn of 2013 and in the spring of 2014, mist nets were put up for capturing birds at the feeding sites in Dæli. Sunflower seeds were provided for a couple of days before the net was put up to familiarise the birds to the site. The birds were transported some distance away from the net in standard cloth bird bags for morphological measuring and behavioural testing. One bird was handled at the time and was released before the next bird was treated. In the present thesis, Andersen's (2012) measures, with some slight adjustments, were used. When captured, the species, sex and age of the birds were recorded. We recorded six measurements of personality, all done by the same person (T. Slagsvold):

1. **Breathing rate.** The number of breaths taken in 30 sec when the bird was held on its back in hand (Figure 1A). Measured by watching chest movements, first right after ring number and colour ID had been recorded, then before releasing the bird after various body measurements were taken some six minutes later. In the subsequent analyses, I used the mean value of the two measurements.
2. **Screaming**. Whether the bird gave none or more distress calls when taken out of the cloth bag.
3. **Biting.** The number of bites when the right index finger was moved towards the bird's bill six times when the bird was held in the left hand (score of 0 - 6) (Figure 1B). This was an improvement from the method used by Andersen (2012) who only used the response variable biting/not biting in a single test.
4. **Tonic immobility**. The bird was held on its back for a couple of seconds until it stopped wriggling. Then it was slowly tilted (in ca. 3 sec) from left to right. The approximate angle the bird had to the palm when it flew away was noted (0°, 60°, 90°, 120°, 150° or 180°) (Figure 1C).

5. **Alarm calling**. Whether the bird screamed while flying away or directly after landing after release from hand, or not.
6. **Flight distance.** The approximate linear distance from the release point from hand and the bird's first landing point (in meters).

Figure 1. In panel A the breathing rate of a great tit is measured, in panel B the biting score is measured, and in panel C the bird's tonic immobility is tested. All performed during the in-hand personality test. Photos: Camilla Thorsteinsen.

2.3.2 Attentiveness test of incubating females

The attentiveness test measured incubating females' persistence to stay on the nest when disturbed by a human. It was performed for the first time in 2014. The box number, date and hour were noted before the trial started. The following was recorded:

1. **Hissing calls.** An index finger was put in the opening of the nest box and wriggled for five seconds. We noted whether the female gave hissing calls or not.
2. **Attentiveness/Persistence to stay on the nest.** We noted if/when the female left the box during the following four behaviours (score of 1-4):
    a. Lid off: The lid of the box was removed without the female bird seeing the observer.
    b. Hand along back side of box: The bird saw the observer putting his/her hand (flat fingers) slowly along the back side of the box down to the nest material for 5 seconds.
    c. Tail touch: The bird's tail was touched for 5 seconds.
    d. Female remains: The focal female bird had still not flown out.
3. **Calling inside.** Whether the female called inside the nest box during any of the four steps above or not.
4. **Distance to perch**. The approximate horizontal distance (in meters) to the first perching site after leaving the nest.
5. **Calling for two min after leaving box**. Whether the female produced warning calls two minutes after she left the box or not.

2.4 Fitness measures

In spring of 2014, the nest boxes were regularly visited (every day/every other day) to follow the developmental stage of the nests and to secure an accurate date of the first egg laid and hatching date for as many nests as possible. The following was recorded: the date of first egg laid, clutch size, duration of incubation period (for last egg laid to first egg hatched), hatching date (for first nestling), number of young hatched, number of fledglings, number of nestlings dying and whether a parent disappeared during the breeding season or not. The percentage mortality of nestlings and the percentage of females that disappeared were calculated in MS Excel (2013). Territories and nests were visited repeatedly and presence or absence of parents recorded. Parents that disappeared during the breeding season after the first egg was laid but before fledging of the brood were assumed to have died.

2.5 Response to a conspecific intruder in the incubation period

A cage containing a conspecific female was placed nearby the test subjects' nest box. The males would get a score of 0-5, where 0 represented that the male never was on top of the cage, and 5 represented that the male was on top of the cage at least one time during each of five consecutive one-minute intervals. For the females, the total time in seconds spent approximately ≤ 2 m from the cage during the five-minute trial (range 0-300 sec) was recorded. This test was performed for the first time during the incubation period in the spring of 2014.

2.6 Response to a predator model in the nestling period

For measuring the birds' response to a predator, a stuffed Tengmalm's owl, *Aegolius funereus*, specimen was used as a predator model. This owl is a natural predator of great tits and other small birds (Haftorn 1971). The test was performed for the first time late in the nestling period (day 10-14) in the spring of 2014. The predator model was placed on the lid of the nest box, and the approximate distance (in meters) between the focal bird(s) and the model during five consecutive one-minute intervals were noted. The mean value of these five measures was used.

2.7 Which responses reflect boldness and shyness?

Merethe Andersen (2012) conducted a master thesis similar to mine of the current population and some of her results will be discussed for help identifying bold and shy responses. She found that repeatability within observers was good but repeatability between observers was poor, and called for better standardisations of the traits. Within the traits, Andersen (2012) found repeatability between the first and second independent measurement (taken in 2011) of screaming and breathing rate for both sexes ($p < 0.05$). Tina Stræte (unpublished data) tested repeatability of the traits measured during the 2014 season. Repeatability of breathing rate was found for both sexes, distance to perch for females (both $p < 0.05$) and attentiveness on the nest for females ($p < 0.10$).

To determine whether the behavioural traits represented reliable methods for assessing personality or not, Andersen (2012) measured the strength of the between-trait relationships. A few significant relationships were found in her study: (1) birds alarm calling when released flew shorter distances from the hand than non-callers. I suggest that birds producing alarm calls (and other sounds) are bold, because the birds risk attracting predators or making themselves more prone to capture by revealing their location (Krams *et al.* 2006). Likewise, I propose that flying shorter compared to longer distances away from a predator may be considered bold, because a shy bird should fly further away to safety and hide out of the predator's sight. (2) Biters screamed more than non-biters. Screaming when trapped by a predator may result in escape by attracting secondary predators or mobbing conspecifics (Møller & Nielsen 2010). Biting may increase the probability of escaping predators (Møller *et al.* 2011) and I assume that biting represents boldness. (3) Screaming birds had higher immobility than non-screaming birds. A high immobility may be considered a bold response, because the bird does not flee at the first chance but waits until completely or almost completely turned around in the hand. Another explanation may be that the shy birds freezes in response to high levels of fear, which will result in high immobility scores (Møller *et al*. 2011; Andersen 2012).

Breathing rate is commonly used as a measure of stress or anxiety (Class *et al.* 2014), and a low breathing rate is therefore considered to be bold. Females with high attentiveness are more persistent to stay on the nest when disturbed by a predator (human) during the incubation period than females with low attentiveness. Highly attentive females are therefore considered more bold than less attentive females. Birds that approached the (caged) conspecific and the predator model closely are also considered bold.

With this information, I have tried to assign the behavioural traits used in the present study to either a bold or shy personality in Table 2.

Table 2. Behavioural traits assigned to either bold or shy personality in great tits based on previous studies results described in the main text.

Behavioural trait	**Bold**	**Shy**
Response when caught by a human:		
Breathing rate	Low	High
Screaming when held	Yes	No
Biting	Yes	No
Tonic immobility	Uncertain	Uncertain
Alarm calling when leaving	Yes	No
Flight distance from hand	Short	Long
Response to human during incubation period:		
Calling inside box	Yes	No
Hissing inside box	Yes	No
Attentiveness on nest	High	Low
Distance from box to perch	Short	Long
Calling 2 min after flying out of box	Yes	No
Response to a conspecific intruder:		
Response to caged bird	Strong	Weak
Response to a predator model:		
Distance to predator model	Short	Long

Note: For descriptions of variables, see main text.

2.8 Statistical methods

The computer software "R" (R Core Team 2014) was used to produce histograms, normal quantile plots (or qq-plots) and Shapiro-Wilk tests for determining normality of the variables. The variables were tested for male values separately, female values separately and the values for both sexes combined. Non-normally distributed values were log transformed or square root transformed in MS Excel (2013). Parametric tests were used when the variables were normally distributed (Student's two sample t-test, Pearson's regression correlation analysis, ANOVA) and non-parametric tests were used when one or more variables were non-normal and could not be log- or square root transformed (Mann-Whitney U test, Spearman rank

correlation test, Pearson's chi-squared test, Fisher's exact test, ANCOVA). Chosen tests are presented in the result tables. Bonferroni corrections were applied when evaluating the final results.

Distance to perching site during the attentiveness test was used as a measure for female flight distance in the analyses, because I believe it is a better predictor of overall fitness (i.e. reproduction and survival) than the flight distance from hand. The latter trait might represent only personal survivability, while the first trait includes abandoning the clutch as well. Hence, the measure of male and female flight distance will originate from different tests.

2.9 Confounding variables

Numerous factors may have influenced personality traits when measured (for instance the date of trial, hour of day or weather conditions) and many variables may have affected measures of fitness in addition to personality traits (for instance territory quality). I have focused on presumably two of the most important confounding variables, namely age of focal bird and time of breeding (egg laying date), in addition to the sex of the bird. The age of the birds might for instance influence their experience, dominance (Dingemanse & de Goede 2004) and territory quality (Perrins & McCleery 1985), and the date of first egg may depend on for instance territory quality, food abundance, temperature and weather (Vedder *et al.* 2013). If fitness measures that significantly differed between female or male age groups (see Appendix table 4) correlated with the date of first egg laid (see Appendix table 6), date of laying was considered a possible confounding factor affecting the behavioural test scores. For males and females, the only measure correlating with date of first egg was hatching date, which was excluded from these analyses due to its close dependency on date of egg laying. Behavioural traits with significant relationships with fitness measures that differed between the age groups (Appendix table 1.1.b, 2.b, and 2.a) were then subjected to analyses where possible effects of the birds' age was accounted for (Appendix table 5).

3 Results

All fitness variables have been compared with all personality traits. I have focused on the results showing significance ($p < 0.05$) or tendencies ($p < 0.10$). Result tables are found in the appendix, but all relationships between personality and reproductive traits are presented in summary tables below (Table 3.a. and 3.b.). Results with $p < 0.10$ are discussed in more detail in the text.

3.1 Response to humans

3.1.1 In-hand personality test

A tendency for earlier egg laying dates was found in females that produced alarm calls compared to females that did not, with an average difference of about two days (Appendix table 1.1.a.). The date of first egg was earlier for slow breathing males than for fast breathing males (Appendix table 2.a.). The clutch size was larger and the hatching date later for males with a long flight distance from hand than for males that flew shorter distances (Appendix table 2.b. and 2.d.). On average, one less nestling was produced by alarm calling females than by non-alarm calling females (Appendix table 1.1.e.). The number of fledglings produced were lower for females that bit more when held in hand than females that bit less (Appendix table 2.f.). However, these relationships were not statistically significant at the 0.05 level.

The number of nestlings dying was twice as many for males that alarm called compared to males that did not (Appendix table 1.2.b., Figure 2A). The percentage mortality of nestlings was larger for males that alarm called than for silent males (Appendix table 1.2.b., Figure 2B). Males with longer flight distance from the hand had clutches that were incubated longer than males with a short flight distance (Appendix table 2.c., Figure 3). These results were significant at the 0.5 level, and the significant relationship between male flight distance from the hand and incubation period held after applying Bonferroni correction for multiple tests.

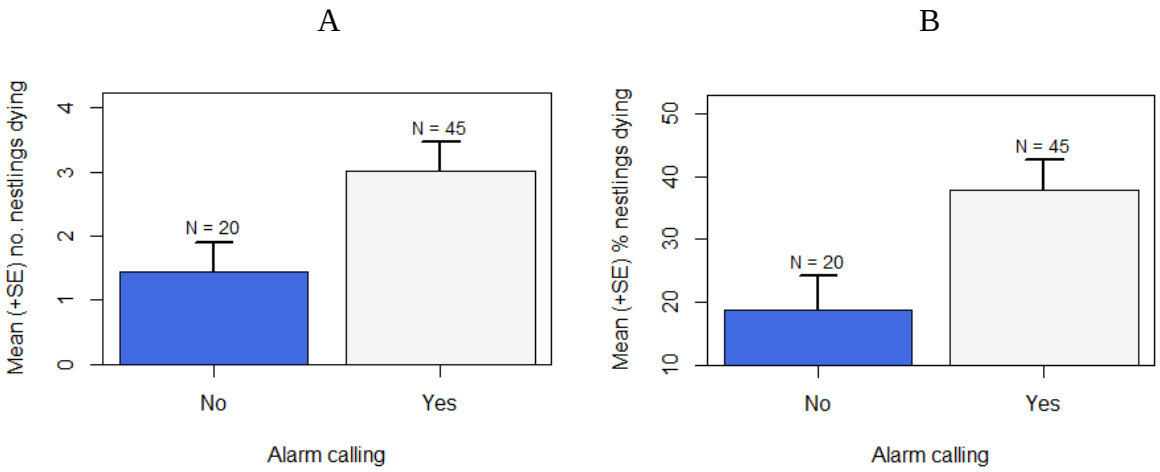

Figure 2. Mean (+SE) number of nestlings dying (A) and mean (+SE) percentage nestlings dying (B) in relation to whether the male great tit parent gave alarm calls or not (A: Z = - 2.20, p = 0.028; B: Z = - 2.37, p = 0.018).

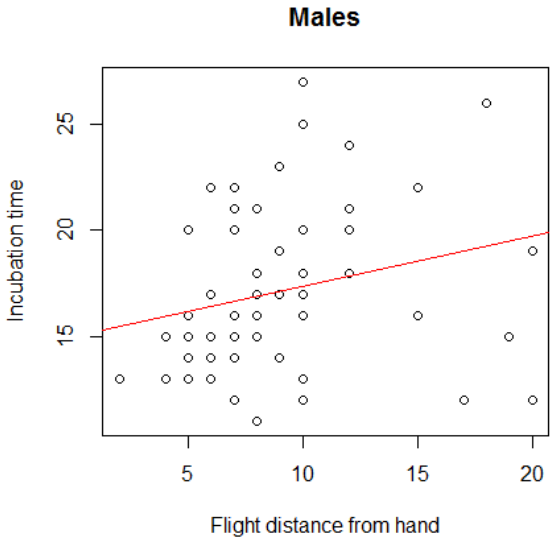

Figure 3. Males with clutches that were incubated for more days had longer flight distances from the hand than males with nests that were incubated less days (r^s = 0.33, N = 68, p = 0.006).

3.1.2 Attentiveness test during the incubation period

The date of first egg laid for females calling for two minutes after exiting the nest box during the attentiveness test were on average six days earlier than for the females that did not call, but the difference was not significant (Appendix table 1.1.a.).

On average, 0.5 more eggs were produced by females that called inside the nest box than by females that did not (Appendix table 1.1.b., Figure 4A). Females staying on the nest

longer when disturbed by a predator (human) had larger clutches than females leaving earlier (Appendix table 2.b., Figure 4B). Hatching dates were on average six days earlier for females calling for two minutes after leaving the box than for non-calling females (Appendix table 1.1.d., Figure 4C). These results were all significant at the 0.05 level, and the relationship between hatching date and calling for two minutes after flight held after applying Bonferroni correction for multiple tests.

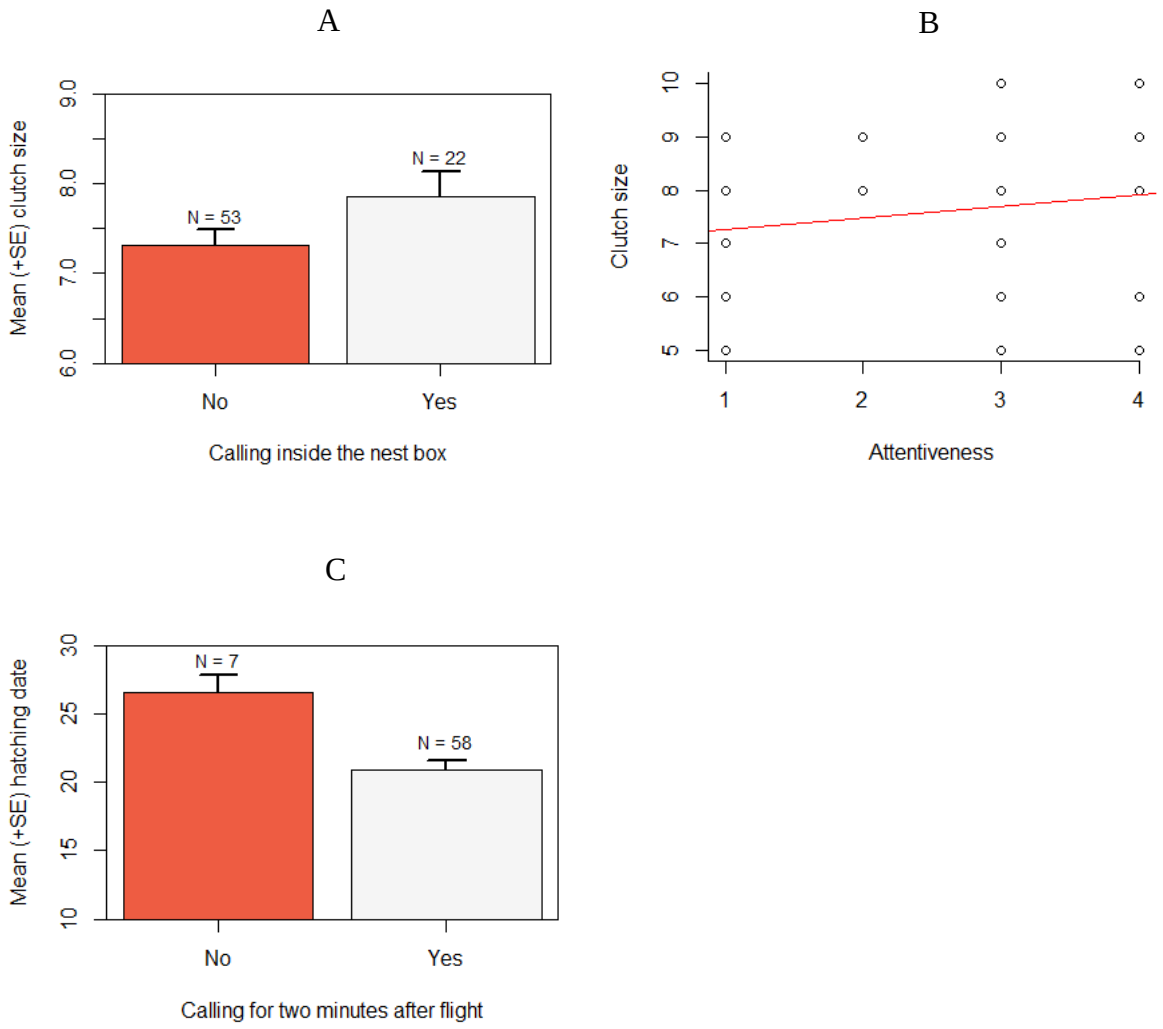

Figure 4. Mean (+SE) clutch size in relation to whether the female great tit parent called inside the nest box when approached by a human or not (A) and to how attentive the female was on the nest when disturbed by a human during the incubation period (B). Mean (+SE) hatching date in relation to whether the female called for two minutes after flying out of the box or not (C) (A: Z = - 2.07, p = 0.039; B: r^s = 0.23, N = 75, p = 0.045; C: t = 2.73, p = 0.008).

Clutch size differed significantly between the two female age groups, with older females having on average 0.6 more eggs than younger females (Appendix table 4.a.). When accounting for the age of female birds the effect of calling inside the nest box on clutch size

was still significant. The same was true for attentiveness. However, significant age effects were found in both tests, meaning that age also influenced clutch size (Appendix table 5).

3.2 Response to disturbances during the incubation and nestling period

Later hatching dates were found for males with strong responses towards the (caged) conspecific female during the incubation period than males not approaching the cage as often, although this result was not significant at the 0.05 level (Appendix table 2.d.).

An earlier date of first egg was found for male and female birds approaching the predator model closely during the nestling period than those who did not (Appendix table 2.a., Figure 5A and B). The hatching date was earlier for close approaching males compared to the more withdrawn males (Appendix table 2.d., Figure 5C). These results were significant at the 0.05 level, and the relationship between date of egg laying and male distance to predator model held after applying Bonferroni correction for multiple tests.

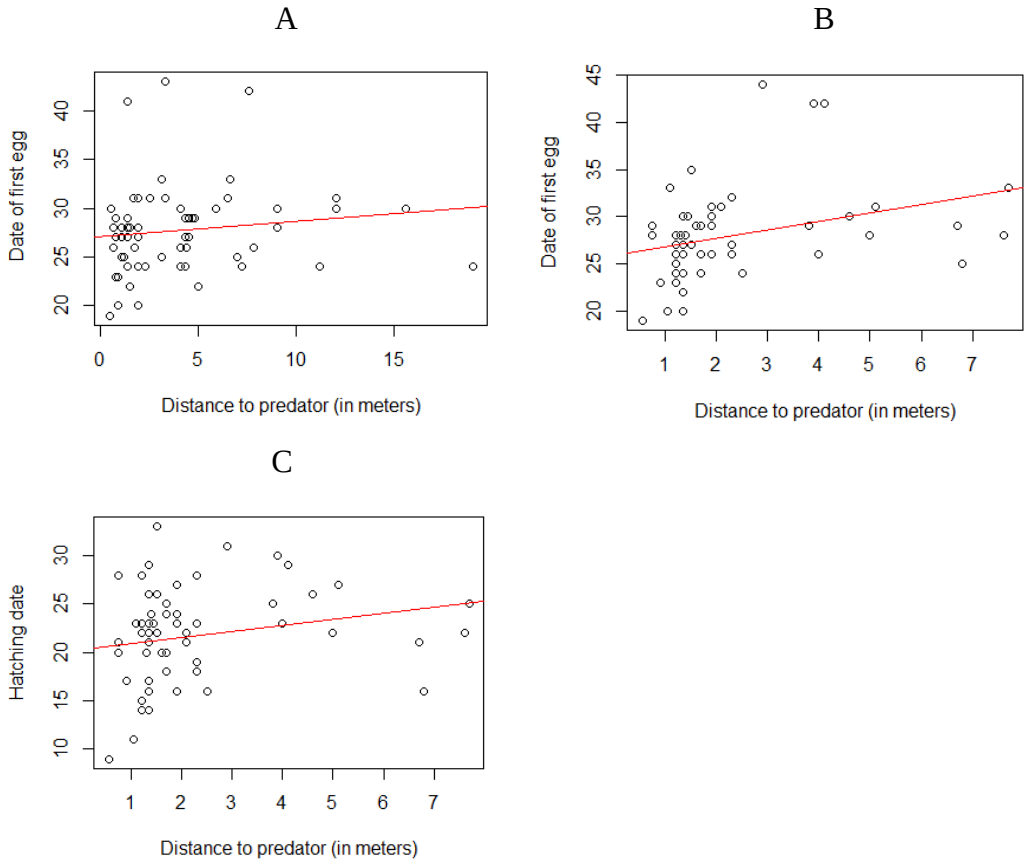

Figure 5. Date of first egg laid in relation to female (A) and male (B) great tit parents' distance to a stuffed owl predator model during the nestling period. Hatching date in relation to male parents distance to predator model (C). (A: $r^s = 0.25$, N = 60, p = 0.05; B: $r^s = 0.42$, N = 55, p = 0.001; C: $r^s = 0.33$, N = 56, p = 0.013)

On average, 1.3 more fledglings were produced by older than by younger males (Appendix table 4.b.). The date of first egg were ca. two days earlier and the hatching date three days earlier for older males, compared to the younger males (Appendix table 4.b.). The date of first egg was still affected by male distance to predator after age was accounted for, but an age effect was also present (Appendix table 5). These results were significant at the 0.05 level.

The number of fledglings in a brood may be a good measure for identifying the parents breeding success because it shows the total number of offspring the parents were able to raise. For that reason, figures illustrating relationships between the number of fledglings and parents behavioural traits are presented (Figure 6A-F and 7A-E) and also for female survival rate (Figure 8A-G).

FEMALES

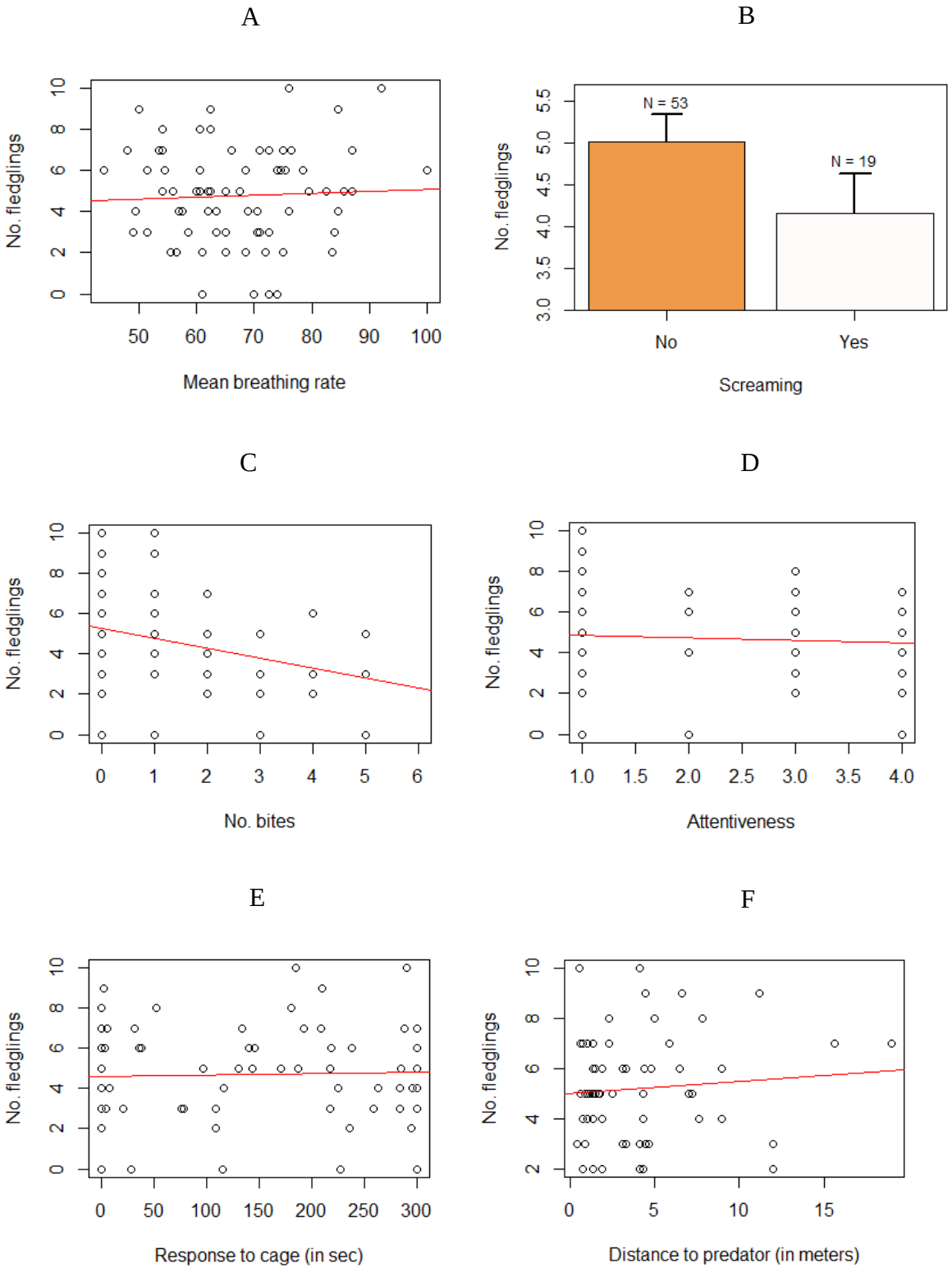

Figure 6. The number of fledglings produced in relation to breathing rate (A), screaming (B), biting (C), attentiveness (D), response to a caged conspecific near the nest during the incubation period (E) and distance to a predator model during the nestling period (F) in female great tit (A: r = 0.04, N = 72, p = 0.71; B: t = 1.38, p = 0.17; C: r^s = -0.22, N = 72, p = 0.061; D: r^s = -0.01, N = 71, p = 0.91; E: r^s = 0.04, N = 71, p = 0.76; F: r^s = 0.07, N = 62, p = 0.59).

MALES

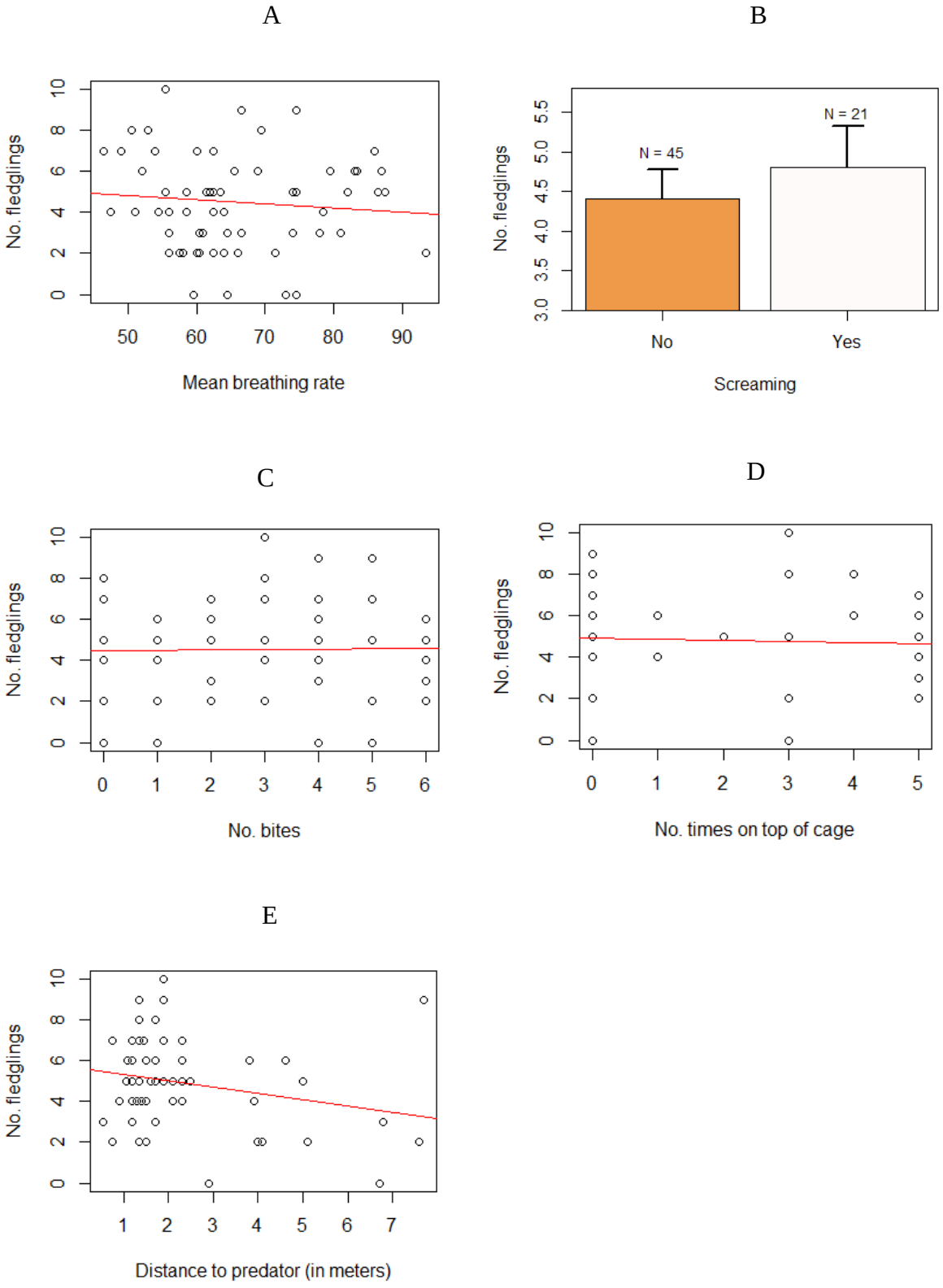

Figure 7. Number of fledglings produced in relation to breathing rate (A), screaming (B), biting (C), response to a caged conspecific near the nest during the incubation period (D) and distance to predator model during the nestling period (E) in male great tits (A: r = -0.12, N = 66, p = 0.35; B: t = 1.38, p = 0.17; C: $r^s < 0.01$, N = 66, p = 0.97; D: $r^s = -0.06$, N = 40, p = 0.70; E: $r^s = -0.13$, N = 56, p = 0.32).

Table 3.a. A summary of the relationships between possible fitness measures and behavioural traits in female great tits.
Relationships with p-values < 0.10 are indicated. For statistical details on relationships, see respective tables in Appendix.

	Fitness measures								
Behavioural traits	Date of first egg	Clutch size	Incubation time	Hatching date	No. nestlings hatching	No. fledglings	No. nestlings dying	% nestling mortality	Female dying
Response to a human:									
Breathing rate[1]	+	+	-	+	+	+	+	+	+
Screaming[2]	-	-	-	-	-	-	+	+	-
Biting[1]	-	-	-	-	+	(-)0.061	+	+	+
Immobility[1]	+	-	-	-	+	+	-	-	+
Alarm calling[2]	(-)0.091	-	+	-	(-)0.060	-	-	-	-
Flight distance[1]	-	+	+	+	-	-	+	+	-
Hissing[2]	+	+	0	+	+	+	+	+	+
Call inside[2]	-	**(+)0.039**	+	+	+	0	+	-	+
Attentiveness[1]	+	(+)**0.045**	+	+	+	-	+	+	-
Call 2min[2]	(-)0.087	+	-	**(-)0.008***	+	0	0	0	-
Response to conspecific:									
Response to cage[1]	-	-	+	-	-	+	-	-	+
Response to predator:									
Distance to predator[1]	(+)**0.050**	-	+	+	+	+	-	-	NA

[1]Correlations with continuous traits (significant or not): +: positive correlation, -: negative correlation, 0: no correlation (p = 1.00).
[2]Comparisons with binary categorical traits (significant or not): +: "Yes" individuals have the highest mean value of the fitness trait, e.g., latest hatching date, -: "Yes" individuals have the lowest mean value of the fitness trait, e.g., earliest hatching date, 0: no difference in means. When compared with female mortality, +: the mortality and personality trait were positively associated, -: the mortality and personality trait were negatively associated. 0: no association.
*The only significant relationship when applying Bonferroni correction for multiple tests.
Note: For descriptions of variables, see main text.

Table 3.b. A summary of the relationships between possible fitness measures and behavioural traits in male great tits.
Relationships with p-values < 0.10 are indicated. For statistical details on relationships, see respective tables in Appendix.

	Fitness measures								
Behavioural traits	Date of first egg	Clutch size	Incubation time	Hatching date	No. nestlings hatching	No. fledglings	No. nestlings dying	% nestling mortality	Male dying
Response to human:									
Breathing rate[1]	(+)0.061	-	+	+	-	-	+	+	+
Screaming[2]	+	+	-	+	+	+	-	-	-
Biting[1]	+	-	-	-	-	+	+	+	-
Immobility[1]	-	+	+	+	+	+	+	+	+
Alarm calling[2]	-	-	+	+	+	-	**(+)0.028**	**(+)0.018**	**+**
Flight distance[1]	-	(+)0.060	**(+)0.006***	(+)0.083	+	+	+	+	+
Response to conspecific:									
Distance to cage[1]	-	-	-	(-)0.096	-	-	-	-	-
Response to predator:									
Distance to predator[1]	**(+)0.001***	-	+	**(+)0.013**	-	-	0	0	NA

[1]Correlations with continuous traits (significant or not): +: positive correlation, -: negative correlation, 0: no correlation (p = 1.00).
[2]Comparisons with binary categorical traits (significant or not): +: "Yes"-individuals have the highest mean value of the fitness trait, e.g., latest hatching date, -: "Yes" individuals have the lowest mean value of the fitness trait, e.g., earliest hatching date, 0: no difference in means. When compared with male mortality, +: the mortality and personality trait are positively associated -: the mortality and personality trait are negatively associated, 0: no association.
*The only significant relationships when applying Bonferroni correction for multiple tests.
Note: For descriptions of variables, see main text.

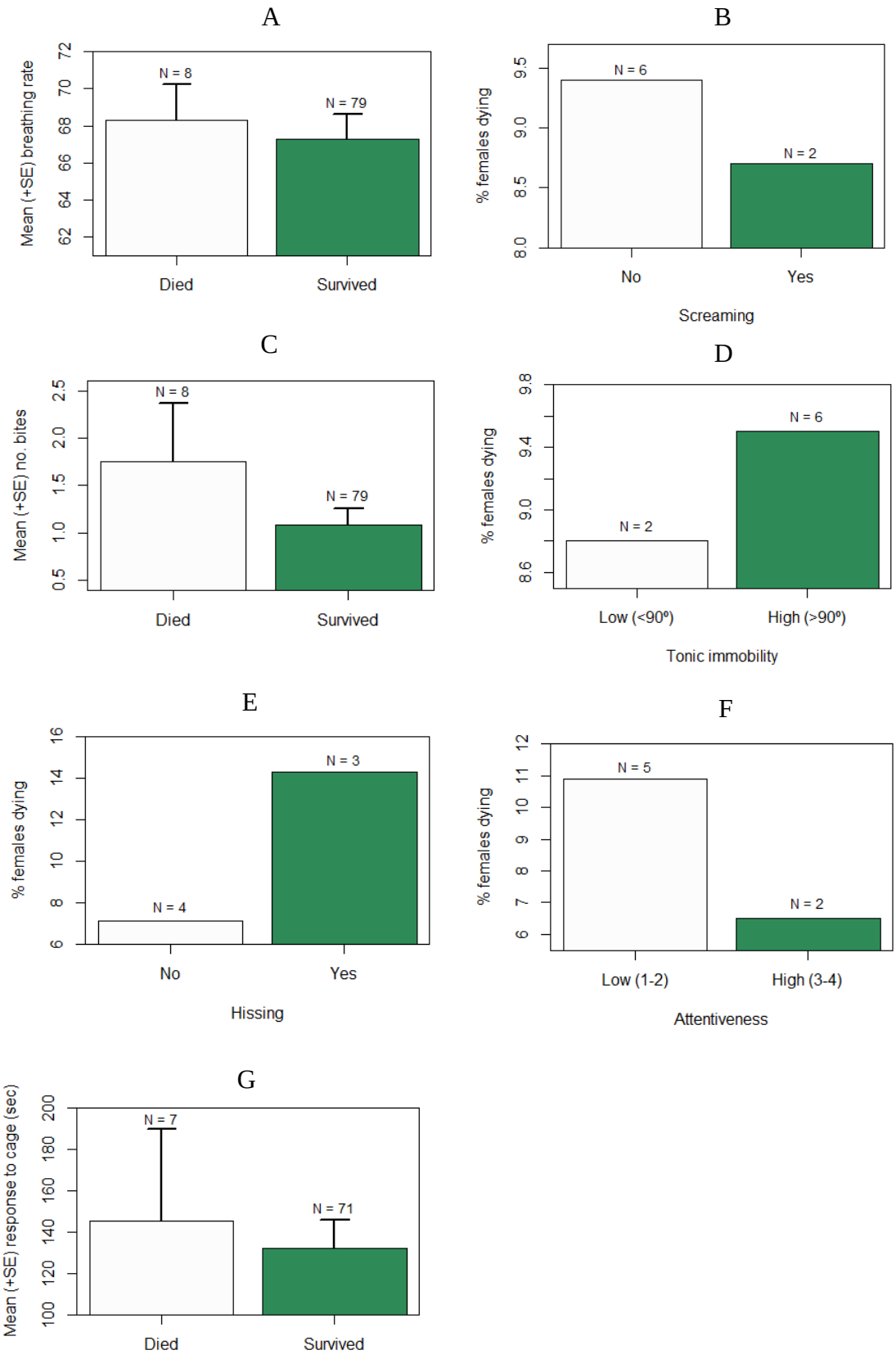

Figure 8. Relationships between mortality during the breeding season and mean (+ SE) breathing rate (A), screaming (B), biting (C), tonic immobility (D), hissing (E), attentiveness (F) and response to (caged) conspecific intruder during the breeding season (G) in female great tits (A: t = 0.24, p = 0.81; B: OR = 0.92, p = 1; C: Z = -1.33, p = 0.18; D: Z = -0.85, p = 0.39; E: OR = 2.14, p = 0.38; F: Z = 0.08, p = 0.94; G: Z = -0.20, p = 0.84).

4 Discussion

In the present study of a great tit population, behavioural traits and fitness variables were measured in the wild and compared between the individuals. From the literature, I found that the available results on relationships between fitness and personality differed substantially, demonstrating a need for more studies on this subject. This was one of the main arguments for conducting this study, which is also the first to obtain results with as many recorded behavioural traits for the same focal bird. The traits I measured cover many aspects of the birds' behaviour: the response towards a predator (human) when caught, attentiveness on the nest during the incubation period with human disturbance, response towards a conspecific intruder during the breeding season, and mobbing behaviour in relation to an avian predator near the brood.

Defining the concepts of boldness and shyness is not straightforward, neither is it easy to determine which traits may represent a bold or a shy personality. Indeed, many studies have tried to determine specific behaviours' placement on the shy-bold axis, i.e. defining which traits may represent boldness or shyness (see for instance Møller *et al.* 2011). Hence, my classification of responses reflecting a bold or shy personality for the various behavioural traits measured in the present study (see Table 2 in the method section), which was based on previous studies results and logic reasoning, may not be correct. However, from hereon I assume that the classification given in Table 2 is correct. Hence, the bold birds will have lower breathing rate, scream, bite more, give alarm calls, fly shorter distances away from the hand and box, and be more aggressive towards conspecifics during the incubation period and predators during the nestling period compared to the shy birds. In addition, bold females will hiss, have high attentiveness scores and call inside and outside of the nest box when disturbed by a human during the incubation period. Support for these descriptions of the bold and shy personalities was found in Stræte's (unpublished) master thesis. She found that great tits with high breathing rates (shy) flew longer distances from the hand and box (shy) and birds that were aggressive towards a caged conspecific (bold) flew shorter distances (bold). She also found that birds alarm calling (bold) had lower immobility scores, indicating that a low immobility might be bold. Her result on biting does not fit my description, with biters (bold) having higher breathing rates (shy) and longer flight distances from hand (shy) than birds that did not bite.

4.1 Reproductive success in relation to personality

Earlier studies on great tit populations show contradictory conclusions (see Table 1), with bold females having higher (Vrublevska *et al.* 2014) or lower reproductive success (Both *et al.* 2005) or with the relationship between breeding success and behaviour varying between years (Dingemanse *et al.* 2004). Results from the present study supporting the prediction that the bold birds should have higher reproductive success than the shy birds were; (1) females and males approaching the predator closely had earlier laying dates (Appendix table 2.a.). (2) Females calling inside the box had larger clutches (Appendix table 1.1.b.). (3) Females with high attentiveness scores had larger clutches (Appendix table 2.b.). (4) Males with shorter flight distances had shorter incubation period (Appendix table 2.c.). (5) Females calling for 2 minutes after leaving the box had earlier hatching dates (Appendix table 1.1.d.), and (6) males approaching the predator model closely had earlier hatching dates (Appendix table 2.d.). All these results were significant at the 0.05 level, but few held when applying a Bonferroni correction (see Table 3.a.-b.). Non-significant results ($p < 0.10$) supporting the same prediction were; (1) females that alarm called had earlier egg laying dates (Appendix table 1.1.a.). (2) Males with low breathing rates had earlier laying dates (Appendix table 2.a.). (3) Females that called for 2 minutes after leaving the box had earlier laying dates (Appendix table 1.1.a.). (4) Males that flew shorter distances from the hand had clutches with earlier hatching dates (Appendix table 2.d), and (5) males that were aggressive against the caged female had earlier hatching dates (Appendix table 2.d.).

Results not supporting the prediction that bold birds have a higher reproductive success than the shy birds were; (1) males that alarm called had more nestlings dying and (2) males that alarm called had higher percentage mortality of nestlings (Appendix table 1.2.b.). This was significant at the 0.05 level, but only one held after Bonferroni correction (see Table 3.b.). Non-significant results were: (1) males with long flight distances had larger clutches (Appendix table 2.b.). (2) Females that alarm called had fewer nestlings (Appendix table 1.1.e.), and (3) biting females had fewer fledglings (Appendix table 2.f.).

In other words, I found some evidence for bold individuals having higher reproductive success, but also for shy individuals having higher success. When considering females only, my results indicate that bold females have higher reproductive success than shy females. Smith and Blumstein's (2007) meta-analysis support that bold animals may have higher reproductive success than shy animals (however, only captive/domestic animals and especially for males). Andersen (2012) found that great tits with large clutch sizes had

significantly lower breathing rates (bold) than birds with small clutch sizes (however, this was not significant after Bonferroni correction). This also supports that bold individuals may have greater reproductive success than shy individuals. In the present study, no relationship between breathing rate and clutch size was found, but a tendency for an earlier date of first egg was found for males with a low breathing rate (Appendix table 2.a.).

Non-significant results ($p < 0.10$) from Andersen's (2012) study were: (1) smaller clutches were laid by screaming (bold) great tits compared to non-screaming birds, suggesting that shy birds had greater reproductive success. 2) Fledglings with highest mean body mass were offspring of birds with low breathing rates (bold). Hence, this indicated that bold birds might have the greatest breeding success. In my study, body measurements were not analysed because of time constraints. 3) More nestlings were produced by birds with low tonic immobility (possibly bold). I found no effect of immobility on either fitness measure.

4.2 Female survival during the breeding season in relation to personality

A relatively high number of females disappeared during the breeding season (~13%), which allowed me to compare the behaviour of birds that disappeared or not. In this study, parent birds that disappeared after the onset of breeding but before fledging of their offspring were considered to have died. No significant differences in behavioural trait scores between females that died or survived were found (Appendix table 3.1.). The absence of relationships between survival and behavioural traits problematizes the determination of which hypothesis presented earlier is most likely. With the mixed evolutionary stable strategies (ESS) hypothesis, I predicted that bold individuals had higher reproductive success but lower survival rate than shy individuals. With the best-of-a-bad-job hypothesis, I predicted that bold individuals had both higher reproductive success and survival than shy individuals. The female birds that disappeared during the breeding season had higher breathing rates (shy), higher biting scores (bold), shorter flight distances (bold) or were more aggressive towards the conspecific during the incubation period (bold) than the surviving females (Appendix table 3.1.a.). The lowest survival rate was found for females not screaming (shy), not alarm calling (shy), hissing (bold), with low attentiveness (shy), calling inside the box (bold) and not calling for two minutes (shy) (Appendix table 3.1.b.). All these results had $p > 0.10$ but may indicate which hypothesis is most likely. These results give equal amounts of support to the ESS prediction that bold birds have the lowest survival, and to the best-of-a-bad-job

prediction that they have the highest survival. The sample size for disappearing males was too small for valid conclusion (but see Appendix table 3.2. for analyses). Smith and Blumstein (2007) found that boldness reduced survival in captive/domestic animals, and especially males.

The studies of Krams *et al*. (2013) and Vrublevska *et al*. (2014) were the only studies I found investigating relationships between survival during the breeding season and personality in wild great tit populations. No significant relationships between female survival during the breeding season and hissing (Appendix table 3.1.b.) or between any of the other fitness measures and hissing (Appendix table 1.1.a-h.) were found in the present thesis. However, Krams *et al*. (2013) found that survival was highest for females that hissed (bold) towards a stuffed great spotted woodpecker, *Dendrocopos major*, predator model. Apparently, predators attacking the females inside the box (feral cats, *Felis silvestris catus*, pine martens, *Martes martes*, weasels, *Mustela nivalis*, stoats, *Mustela erminea*, and woodpeckers) caused most of the mortality. In our study, predators attacking outside of the nest box, e.g., European sparrowhawk, *Accipiter nisus*, probably caused most of the adult mortality. Krams *et al*. (2013) found no relationships between clutch size (i.e., reproductive success) and hissing. Environmental or other confounding factors may affect the proportion of hissing females in a population. For instance, in Krams *et al*. (2013) study, the number of hissing compared to non-hissing females increased following severe, cold winters. The authors suggested this to be a consequence of more bold, dominant females surviving the winter than less dominant females. In the present study, hissing was tested only during a single season (2014). The type of predator presented during the measuring of the trait may also influence the number of females hissing. In the present study, the predator was a human that was not seen approaching (but possibly heard) before placing a finger in the opening of the box, and approximately 1/4 of the females hissed. In Krams *et al*. (2013) study, the predator was a stuffed woodpecker model placed in the opening of the box and more than 2/3 of the females hissed. Hence, the effectiveness of hissing as an antipredator and/or nest defence behaviour is still unclear. However, Krams *et al*. (2013) results support the best-of-a-bad-job prediction that bold individuals have higher survival, although no relationship between reproduction (clutch size) and hissing was found.

In Vrublevska *et al*.'s (2014) study of great tits, fast-exploring females (tested in captivity) had higher survival during the breeding season and higher nest success than slow-exploring females. Explorative behaviour has not been tested for the current population.

However, because exploration is highly associated with boldness (Wilson *et al.* 1993), Vrublevska *et al*.'s (2014) findings are interesting. A common assumption is that fast exploring individuals are bold, while slow exploring individuals are shy (see for instance Cole & Quinn 2014). Vrublevska *et al.*'s (2014) findings hence indicate that higher survival and greater reproductive success are related to boldness, supporting the best-of-a-bad job hypothesis where boldness is the superior phenotype.

Andersen (2012) found that birds surviving the winter screamed less (shy) than birds that did not survive (however, this was not significant after Bonferroni correction). This supports the ESS prediction saying that boldness should reduce survival. I did not study survival from one year to the next and found no relationship between female survival during the breeding season and screaming (Appendix table 3.a.-b.). Using data on survival from more than one breeding season is probably important for determining whether the bold birds have the lowest survival or not, i.e., determining which hypothesis is most likely. For instance, fitness consequences of different personalities may vary between seasons or from one year to the next, as demonstrated for instance in Dingemanse *et al.*'s (2004) study on great tits where the relationship between exploration and survival differed between years and sexes.

4.3 Limitations of present study and further work

It needs to be considered that the traits we measured in the field might not represent boldness or shyness, or that they should be measured differently for valid conclusions. The absence of relationships between observed behavioural traits and fitness may have simply arisen from problems with the measuring procedures. Even more trials per individual may be important. We could also have taken the birds into the lab to do the exploration test for comparison between this "classical" personality measure and the traits we collected in the field. Testing for repeatability of behavioural traits may indicate whether they are good indicators of personality or not. Andersen (2012) and Stræte (unpublished data) tested repeatability of the behavioural traits measured on the current population in the 2011 and 2014 season, respectively. Both studies described low repeatability of the traits. Although a few were statistically significant, the amount of variation explained (r^2) was often low. As requested by Andersen (2012), better standardisations of traits should therefore be strived for.

The data used in this study have been collected only from one breeding season. However, the sample size from this season was quite large (199 birds) and I compared my

results (2014) with those from Andersen's (2012) study collected in 2011 on the current population. In 2014, the mean clutch size was below average, and the number of fledglings were below the long time average (T. Slagsvold, unpublished data). Because the breeding success was low, we would have expected effects of personality differences on the reproduction and survival of the birds.

Many confounding variables were probably present in this study, and I chose to investigate only two variables likely to have affected the results most strongly, namely age and laying date (i.e., breeding time), in addition to sex. Significant effects were found for age, but not for breeding time. The results from the personality tests were still significant when age was accounted for (Appendix table 5). It may be necessary to account for more confounding factors in further studies, for instance weather conditions, date, time of day or condition of the bird. Earlier studies on fitness and behaviour in great tit populations (see Table 1) accounted for sex, age and year but few or no other confounding factors.

4.4 Conclusion

A great tit population was studied in the wild in a breeding season (2014) with relatively low reproductive success, thus under conditions when fitness differences in relation to personality were expected. I have measured more traits in the wild than previous studies, and these covered a range of conditions, namely response to a human and an avian predator and a conspecific (caged) intruder. I also compared survival rate of female great tits during the breeding season with a number of different personality traits, which has never been analysed previously.

With the ESS hypothesis, I predicted that bold birds had greater reproductive success but lower survival rate than shy birds. With the alternative best-of-a-bad-job hypothesis, I predicted that bold birds both had greater reproduction and higher survival than shy birds. In this study, no clear relationship existed between fitness and personality. My results indicated that survival during the breeding season was equal for the shy and bold birds (however, only non-significant results ($p > 0.10$)) and that bold birds had higher reproductive success than shy birds, especially for females. However, some results indicated that shy individuals had higher breeding success than bold individuals. Bonferroni corrections were applied to the final results and resulted in only three p-values remaining significant. These results supported the prediction that bold birds have the greatest reproductive success (see Table 3.a.-b.).

Andersen (2012) found that winter survival might be higher for shy birds in the current population. Together, these results may indicate a trade-off between reproduction and survival outside of the breeding season, in support of the mixed ESS hypothesis.

However, Krams *et al.* (2013) found that bold (hissing) female great tits had the highest survival during the breeding season and Vrublevska *et al.* (2014) found that bold (fast exploring) female great tits had higher survival during the breeding season and nests success than shy females. Hence, data from three different studies indicate that bold great tit females may have the highest reproductive success and survival during the breeding season, in support of the best-of-a-bad-job hypothesis.

Further studies should test repeatability of behavioural traits for assessing whether they are good indicators of personality or not, and develop and/or use standardised traits and tests for different focal species. More trials per individual may be important for determining the importance (or lack thereof) of the measured behaviour in contexts regarding reproductive success and survival. I found no relationship between female survival during the breeding season and personality, which may have been caused for instance by too few study seasons. Behaviour and fitness variables should therefore be tested over several years, as effects may vary between years or seasons. I accounted for two confounding variables, in addition to sex, namely breeding time (date of first egg) and age. More confounding variables should be tested, for instance weather conditions and time of day.

References

Andersen, M., 2012. *What is maintaining variation in personalities in wild populations of great tits and blue tits?* Master thesis. Oslo: Department of Biosciences, University of Oslo.

Bell, A.M., Hankison, S.J. & Laskowski, K.L., 2009. The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, **77**: 771–783.

Boon, A.K., Réale, D. & Boutin, S., 2007. The interaction between personality, offspring fitness and food abundance in North American red squirrels. *Ecology Letters*, **10**: 1094–1104.

Both, C. *et al.*, 2005. Pairs of extreme avian personalities have highest reproductive success. *Journal of Animal Ecology*, **74**: 667–674.

Boulton, K. *et al.*, 2014. How stable are personalities? A multivariate view of behavioural variation over long and short timescales in the sheepshead swordtail, *Xiphophorus birchmanni*. *Behavioral Ecology and Sociobiology*, **68**: 791–803.

Class, B., Kluen, E. & Brommer, J.E., 2014. Evolutionary quantitative genetics of behavioral responses to handling in a wild passerine. *Ecology and Evolution*, **4**: 427–440.

Cole, E.F. & Quinn, J.L., 2014. Shy birds play it safe: personality in captivity predicts risk responsiveness during reproduction in the wild. *Biology Letters*, **10**: 20140178. doi: 10.1098/rsbl.2014.0178

Coleman, K. & Wilson, D.S., 1998. Shyness and boldness in pumpkinseed sunfish: individual differences are context-specific. *Animal Behaviour*, **56**: 927–936.

Colléter, M. & Brown, C., 2011. Personality traits predict hierarchy rank in male rainbowfish social groups. *Animal Behaviour*, **81**: 1231–1237.

Dingemanse, N.J. *et al.*, 2004. Fitness consequences of avian personalities in a fluctuating environment. *Proceedings of the Royal Society B*, **271**: 847–852.

Dingemanse, N.J. *et al.*, 2002. Repeatability and heritability of exploratory behaviour in great tits from the wild. *Animal Behaviour*, **64**: 929–938.

Dingemanse, N.J. & de Goede, P., 2004. The relation between dominance and exploratory behavior is context-dependent in wild great tits. *Behavioral Ecology*, **15**: 1023–1030.

Gosling, S.D., 2001. From mice to men: what can we learn about personality from animal research? *Psychological Bulletin*, **127**: 45–86.

Gosling, S.D. & Vazire, S., 2002. Are we barking up the right tree? Evaluating a comparative approach to personality. *Journal of Research in Personality*, **36**: 607–614.

Groothuis, T.G.G. & Carere, C., 2005. Avian personalities: characterization and epigenesis. *Neuroscience and Biobehavioral Reviews*, **29**: 137–150.

Haftorn, S., 1971. *Norges fugler*, Trondheim: Universitetsforlaget.

Hansen, B.T., Johannessen, L.E. & Slagsvold, T., 2007. No cultural transmission of species recognition between parents and offspring in free-living great tits and blue tits. *Behavioral Ecology and Sociobiology*, **61**: 1203–1209.

Krams, I. *et al.*, 2013. Hissing calls improve survival in incubating female great tits (*Parus major*). *Acta Ethologica*, **17**: 83–88.

Krams, I., Krama, T. & Igaune, K., 2006. Alarm calls of wintering great tits *Parus major*: warning of mate, reciprocal altruism or a message to the predator? *Journal of Avian Biology*, **37**: 131–136.

Malmkvist, J. & Hansen, S.W., 2002. Generalization of fear in farm mink, *Mustela vison*, genetically selected for behaviour towards humans. *Animal Behaviour*, **64**: 487–501.

Møller, A.P., Christiansen, S.S. & Mousseau, T.A., 2011. Sexual signals, risk of predation and escape behavior. *Behavioral Ecology*, **22**: 800–807.

Møller, A.P. & Nielsen, J.T., 2010. Fear screams and adaptation to avoid imminent death: effects of genetic variation and predation. *Ethology Ecology & Evolution*, **22**: 183–202.

van Oers, K. *et al.*, 2004. Realized heritability and repeatability of risk-taking behaviour in relation to avian personalities. *Proceedings of the Royal Society B*, **271**: 65–73.

Perrins, C.M. & McCleery, R.H., 1985. The effect of age and pair bond on the breeding success of great tits *Parus major*. *Ibis*, **127**: 306–315.

R Core Team, 2014. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. URL http://www.r-project.org/.

Sih *et al.*, 2004a. Behavioural syndromes: an integrative overview. *The Quartlerly Review of Biology*, **79**: 241–277.

Sih, A., Bell, A. & Johnson, J.C., 2004b. Behavioral syndromes: an ecological and evolutionary overview. *Trends in Ecology & Evolution*, **19**: 372–378.

Slagsvold, T. *et al.*, 2002. Mate choice and imprinting in birds studied by cross-fostering in the wild. *Proceedings of the Royal Society B* , **269**: 1449–1455.

Slagsvold, T. *et al.*, 2013. Vertical and horizontal transmission of nest site preferences in titmice. *Animal Behaviour*, **85**: 323–328.

Slagsvold, T. & Wiebe, K.L., 2007. Learning the ecological niche. *Proceedings of the Royal Society B*, **274**: 19–23.

Smith, B.R. & Blumstein, D.T., 2007. Fitness consequences of personality: a meta-analysis. *Behavioral Ecology*, **19**: 448–455.

Stearns, S.C., 1992. Evolutionary explanation. In *The evolution of life histories*. New York, US: Oxford University Press Inc., pp. 9–18.

Svensson, L., 1975. *Identification guide to european passerines* 2nd ed., Stocholm: Naturhistoriska Riksmuseet.

Vedder, O., Bouwhuis, S. & Sheldon, B.C., 2013. Quantitative assessment of the importance of phenotypic plasticity in adaptation to climate change in wild bird populations. *PLoS Biology*, **11**: e1001605.

Vrublevska, J. *et al.*, 2014. Personality and density affect nest defence and nest survival in the great tit. *Acta Ethologica*, **17**: 1–10.

Wilson, D.S. *et al.*, 1993. Shy-bold continuum in pumpkinseed sunfish (*Lepomis gibbosus*): an ecological study of a psychological trait. *Journal of Comparative Psychology*, **107**: 250–260.

Wilson, D.S. *et al.*, 1994. Shyness and boldness in humans and other animals. *Trends in Ecology & Evolution*, **9**: 442–446.

Appendix

Appendix table 1.1.:

Five binary behavioural traits in relation to eight fitness variables in female great tits.

Table a.	**Date of first egg (1 = 1 April)**			
	No	**Yes**		
Behavioural trait	**mean ± SD (N)**	**mean ± SD (N)**	**t/z**	**p**
Screaming[a]	28.0 ± 5.6 (58)	27.6 ± 3.9 (22)	0.42[2]	0.68
Alarm calling[a]	29.3 ± 5.2 (23)	27.4 ± 5.1 (57)	1.70[2]	**0.091**
Hissing[b]	28.1 ± 5.9 (55)	28.5 ± 4.1 (20)	- 0.69[2]	0.50
Call inside nest box[b]	28.3 ± 6.0 (53)	27.7 ± 3.8 (22)	0.27[2]	0.79
Call 2 min after flight[b]	33.6 ± 8.2 (7)	27.5 ± 5.2 (56)	1.72[2]	**0.087**
Table b.		**Clutch size**		
Screaming[a]	7.46 ± 1.36 (57)	7.45 ± 1.32 (20)	0.17[2]	0.87
Alarm calling[a]	7.52 ± 1.29 (21)	7.43 ± 1.37 (56)	0.25[2]	0.80
Hissing[b]	7.45 ± 1.29 (55)	7.55 ± 1.19 (20)	- 0.25[2]	0.81
Call inside nest box[b]	7.32 ± 1.21 (53)	7.86 ± 1.32 (22)	- 2.07[2]	**0.039**
Call 2 min after flight[b]	7.00 ± 1.29 (7)	7.50 ± 1.18 (56)	- 1.07[2]	0.29
Table c.		**Duration of incubation period**		
Screaming[a]	17.1 ± 3.9 (53)	16.9 ± 3.0 (20)	-0.18[2]	0.86
Alarm calling[a]	17.0 ± 3.7 (21)	17.1 ± 3.6 (52)	- 0.22[2]	0.83
Hissing[b]	17.0 ± 3.9 (55)	17.0 ± 2.9 (20)	- 0.11[2]	0.91
Call inside nest box[b]	16.9 ± 3.9 (53)	17.2 ± 3.0 (22)	- 0.93[2]	0.35
Call 2 min after flight[b]	17.0 ± 4.7 (7)	16.8 ± 3.4 (56)	0.03[2]	0.97
Table d.		**Hatching date (1 = 1 May)**		
Screaming[a]	21.4 ± 5.6 (57)	21.1 ± 4.6 (20)	0.17[1]	0.86
Alarm calling[a]	22.9 ± 6.0 (22)	20.6 ± 5.4 (55)	1.63[1]	0.11
Hissing[b]	21.5 ± 5.9 (56)	22.2 ± 4.0 (22)	0.48[1]	0.63
Call inside nest box[b]	21. 7 ± 6.0 (55)	21.9 ± 3.6 (23)	0.11[1]	0.91
Call 2 min after flight[b]	26.6 ± 3.3 (7)	20.9 ± 5.3 (58)	2.73[1]	**0.008***
Table e.		**Number of nestlings hatching**		
Screaming[a]	7.29 ± 2.14 (56)	6.90 ± 2.22 (20)	0.69[1]	0.50
Alarm calling[a]	7.91 ± 2.43 (22)	6.89 ± 1.98 (54)	1.91[1]	**0.060**
Hissing[b]	7.14 ± 2.14 (56)	7.82 ± 2.02 (22)	- 1.11[2]	0.27
Call inside nest box[b]	7.22 ± 2.14 (55)	7.61 ± 2.08 (23)	- 0.73[2]	0.46

Call 2 min after flight[b]	6.71 ± 2.21 (7)	7.47 ± 1.95 (58)	0.95[1]	0.35

Table f.	**Number of fledglings**			
Screaming[a]	5.0 ± 2.4 (53)	4.2 ± 2.0 (19)	1.38[1]	0.17
Alarm calling[a]	4.9 ± 2.7 (22)	4.8 ± 2.2 (50)	0.17[1]	0.86
Hissing[b]	4.4 ± 2.6 (21)	4.9 ± 2.4 (50)	0.75[1]	0.46
Call inside nest box[b]	4.7 ± 2.3 (49)	4.7 ± 2.8 (22)	0.02[1]	0.98
Call 2 min after flight[b]	4.6 ± 3.1 (5)	4.6 ± 2.4 (53)	0.14[2]	0.89

Table g.	**Number of nestlings dying**			
Screaming[a]	2.4 ± 2.8 (52)	2.6 ± 2.5 (19)	- 0.55[2]	0.58
Alarm calling[a]	3.0 ± 3.4 (22)	2.1 ± 2.4 (49)	0.86[2]	0.39
Hissing[b]	2.3 ± 2.6 (50)	3.3 ± 3.5 (21)	- 1.22[2]	0.22
Call inside nest box[b]	2.6 ± 2.6 (49)	2.8 ± 3.6 (22)	0.39[2]	0.70
Call 2 min after flight[b]	2.8 ± 3.1 (5)	2.8 ± 2.6 (53)	- 0.07[2]	0.94

Table h.	**% mortality of nestlings**			
Screaming[a]	28.9 ± 30.4 (52)	33.6 ± 31.1 (19)	- 0.60[2]	0.55
Alarm calling[a]	34.7 ± 33.2 (22)	28.1 ± 29.2 (49)	0.73[2]	0.46
Hissing[b]	29.5 ± 30.1 (50)	38.7 ± 35.2 (21)	- 1.10[2]	0.27
Call inside nest box[b]	32.3 ± 30.4 (49)	32.2 ± 36.9 (22)	0.42[2]	0.68
Call 2 min after flight[b]	36.0 ± 37.9 (5)	36.0 + 31.1 (53)	- 0.07[2]	0.94

[1]Student's two sample t-test
[2]Mann-Whitney U test
*Significant after Bonferroni correction
[a]Measured during the in-hand personality test
[b]Measured during the attentiveness test
Note: For descriptions of variables, see main text.

Appendix table 1.2.:

Testing fitness variables in relation to screaming and alarm calling in male great tits. Screaming and alarm calling were measured during the in-hand personality test.

Table a.	**Screaming**			
	No	**Yes**		
Variable	**Mean ± SD (N)**	**Mean ± SD (N)**	**Z/t**	**p**
Date of first egg (1 = 1 April)	27.5 ± 4.3 (51)	28.4 ± 5.2 (25)	0.08[2]	0.94
Clutch size	7.3 ± 1.4 (49)	7.4 ± 1.3 (23)	0.06[2]	0.95
Duration of incubation period	17.2 ± 3.5 (45)	16.9 ± 3.8 (23)	0.47[2]	0.64
Hatching date (1 = 1 May)	20.7 ± 5.4 (48)	21.7 ± 5.6 (23)	0.76[1]	0.45
No. nestlings hatching	7.1 ± 2.3 (47)	7.2 ± 2.4 (23)	0.19[1]	0.85
No. fledglings	4.4 ± 2.5 (45)	4.8 ± 2.3 (21)	0.64[1]	0.53
No. nestlings dying	2.7 ± 3.1 (44)	2.2 ± 2.2 (21)	0[2]	1
% mortality of nestlings	33.1 ± 34.1 (44)	29.5 ± 27.0 (21)	0.02[2]	0.98

Table b.	**Alarm calling**			
Variable	**No**	**Yes**		
Date of first egg (1 = 1 April)	27.9 ± 4.2 (26)	27.7 ± 4.8 (50)	0.35[2]	0.73
Clutch size	7.5 ± 1.6 (25)	7.3 ± 1.2 (47)	0.67[2]	0.51
Duration of incubation period	16.8 ± 3.4 (24)	17.2 ± 3.7 (44)	-0.30[2]	0.77
Hatching date (1 = 1 May)	20.8 ± 5.8 (25)	21.2 ± 5.3 (46)	0.26[1]	0.80
No. nestlings hatching	6.9 ± 2.5 (24)	7.3 ± 2.2 (46)	0.70[1]	0.49
No. fledglings	5.0 ± 2.0 (21)	4.3 ± 2.6 (45)	1.08[1]	0.28
No. nestlings dying	1.5 ± 2.0 (20)	3.0 ± 3.0 (45)	-2.20[2]	**0.028**
% mortality of nestlings	18.7 ± 25.1 (20)	37.8 ± 32.9 (45)	-2.37[2]	**0.018**

[1]Students two sample t-test
[2]Mann-Whitney U test
Note: For descriptions of variables, see main text.

Appendix table 2:

Correlations between seven behavioural traits and nine fitness variables in female and male great tits. All values have been treated as numeric values.

Table a.				**Date of first egg**		
	Females			**Males**		
Behavioural trait	**r/rs**	**N**	**p**	**r/rs**	**N**	**p**
Response to human:						
Breathing rate[a]	0.09[2]	80	0.41	0.22[2]	76	**0.061**
Biting[a]	- 0.05[2]	80	0.66	0.08[2]	76	0.48
Immobility[a]	0.11[2]	80	0.34	- 0.10[2]	76	0.40
Attentiveness[b]	0.15[2]	75	0.20			
Flight distance[abc]	- 0.02[2]	67	0.88	- 0.13[2]	76	0.27
Response to conspecific:						
Response to cage	- 0.13[2]	76	0.26	- 0.17[2]	43	0.29
Response to predator:						
Distance to predator	0.25[2]	60	**0.050**	0.42[2]	55	**0.001***

Table b.				**Clutch size**		
Response to human:						
Breathing rate[a]	0.06[2]	77	0.61	- 0.18[2]	72	0.14
Biting[a]	- 0.05[2]	77	0.67	- 0.01[2]	72	0.91
Immobility[a]	- 0.11[2]	77	0.31	0.00[2]	72	0.98
Attentiveness[b]	0.23[2]	75	**0.045**			
Flight distance[abc]	0.01[2]	67	0.93	0.22[2]	72	**0.060**
Response to conspecific:						
Response to cage	- 0.07[2]	76	0.56	- 0.12[2]	43	0.45
Response to predator:						
Distance to predator	- 0.06[2]	60	0.62	- 0.22[2]	55	0.11

Table c.				**Duration of incubation period**		
Response to human:						
Breathing rate[a]	- 0.04[2]	73	0.77	0.12[2]	68	0.32
Biting[a]	- 0.03[2]	73	0.80	- 0.12[2]	68	0.31
Immobility[a]	- 0.18[2]	73	0.13	0.06[2]	68	0.64
Attentiveness[b]	0.13[2]	75	0.27			
Flight distance[abc]	0.02[2]	67	0.90	0.33[2]	68	**0.006***
Response to conspecific:						
Response to cage	0.02[2]	75	0.88	- 0.09[2]	43	0.55
Response to predator:						
Distance to predator	0.01[2]	60	0.95	0.05[2]	55	0.73

Table d.	**Hatching date**
Response to human:	

Breathing rate[a]	0.11[1]	77	0.32	0.18[1]	71	0.14
Biting[a]	- 0.12[2]	77	0.29	- 0.05[2]	71	0.71
Immobility[a]	- 0.03[2]	77	0.82	0.01[2]	71	0.94
Attentiveness[b]	0.17[2]	78	0.15			
Flight distance[abc]	0.13[1]	70	0.28	0.21[2]	71	**0.083**
Response to conspecific:						
Response to cage	- 0.10[2]	78	0.38	- 0.25[2]	45	**0.096**
Response to predator:						
Distance to predator	0.20[1]	62	0.11	0.33[2]	56	**0.013**

Table e.			**Number of nestlings hatching**			
Response to human:						
Breathing rate[a]	0.15[1]	76	0.20	- 0.02[1]	70	0.85
Biting[a]	0.03[2]	76	0.77	- 0.01[2]	70	0.95
Immobility[a]	0.07[2]	76	0.57	0.18[2]	70	0.13
Attentiveness[b]	0.02[2]	78	0.88			
Flight distance[abc]	- 0.09[1]	70	0.47	0.08[2]	70	0.50
Response to conspecific:						
Response to cage	- 0.07[2]	78	0.53	- 0.03[2]	45	0.83
Response to predator:						
Distance to predator	0.09[2]	62	0.50	- 0.11[2]	56	0.41

Table f.			**Number of fledglings**			
Response to human:						
Breathing rate[a]	0.04[1]	72	0.71	- 0.12[1]	66	0.35
Biting[a]	- 0.22[2]	72	**0.061**	0.00[2]	66	0.97
Immobility[a]	0.06[2]	72	0.63	0.10[2]	66	0.44
Attentiveness[b]	- 0.01[2]	71	0.91			
Flight distance[abc]	- 0.11[2]	63	0.38	0.02[2]	66	0.86
Response to conspecific:						
Response to cage	0.04[2]	71	0.76	- 0.06[2]	40	0.70
Response to predator:						
Distance to predator	0.07[2]	62	0.59	- 0.13[2]	56	0.32

Table g.			**Number of nestlings dying**			
Response to human:						
Breathing rate[a]	0.11[2]	71	0.37	0.04[2]	65	0.73
Biting[a]	0.11[2]	71	0.38	0.04[2]	65	0.75
Immobility[a]	- 0.02[2]	71	0.87	0.11[2]	65	0.38
Attentiveness[b]	0.03[2]	71	0.80			
Flight distance[abc]	0.06[2]	63	0.67	0.08[2]	65	0.52
Response to conspecific:						
Response to cage	- 0.12[2]	71	0.31	- 0.02[2]	40	0.91
Response to predator:						
Distance to predator	- 0.04[2]	62	0.75	- 0.00[2]	56	0.97

Table h.				**% mortality of nestlings**		
Response to human:						
Breathing rate[a]	0.08[2]	71	0.50	0.05[2]	65	0.67
Biting[a]	0.10[2]	71	0.42	0.03[2]	65	0.79
Immobility[a]	- 0.03[2]	71	0.83	0.11[2]	65	0.39
Attentiveness[b]	0.02[2]	71	0.90			
Flight distance[abc]	0.10[2]	63	0.42	0.05[2]	65	0.69
Response to conspecific:						
Response to cage	- 0.11[2]	71	0.37	- 0.02[2]	40	0.91
Response to predator:						
Distance to predator	- 0.04[2]	62	0.75	0.00[2]	56	0.99

Table i.				**Female/male disappeared before fledging**		
Response to human:						
Breathing rate[a]	0.05[2]	87	0.66	0.05[2]	81	0.63
Biting[a]	0.14[2]	87	0.18	- 0.07[2]	81	0.52
Immobility[a]	0.09[2]	87	0.40	0.16[2]	81	0.15
Attentiveness[b]	- 0.01[2]	77	0.94			
Flight distance[abc]	- 0.04[2]	69	0.76	0.03[2]	81	0.77
Response to conspecific:						
Response to cage	0.02[2]	78	0.84	- 0.14[2]	45	0.35
Response to predator:						
Distance to predator	-	-	-	-	-	-

[1]Pearson's regression correlation test
[2]Spearman rank correlation test
[a]Measured during the in-hand personality test
[b]Measured during the attentiveness test
[c]Different measure for males and females
*Significant after Bonferroni correction
Note: For descriptions of variables, see main text.

Appendix table 3.1.:

Table a. Comparing five continuous behavioural traits between female great tits that died or survived during the breeding season.

	Died	**Survived**		
Behavioural trait	**mean ± SD (N)**	**mean ± SD (N)**	**t/Z**	**p**
Breathing rate[a]	68.3 ± 5.4 (8)	67.3 ± 11.8 (79)	0.24[1]	0.81
Biting[a]	1.75 ± 1.75 (8)	1.09 ± 1.47 (79)	- 1.33[2]	0.18
Flight distance[b]	7.67 ± 2.94 (6)	8.78 ± 4.5 (63)	0.31[2]	0.76
Response to cage	145.4 ± 116.9 (7)	132.04 ± 117.25 (71)	- 0.20[2]	0.84
Distance to predator[c]	(0)	4.13 ± 3.94 (62)	-	-

[1]Student's two sample t-test
[2]Mann-Whitney U test
[a]Measured during the in-hand personality test
[b]Measured during the attentiveness test
[c]None of the tested individuals died.
Note: For description of the variables, see main text.

Table b. Comparing seven categorical behavioural traits between female great tits that died or survived during the breeding season.

Behavioural trait		Survived	Died	Z/OR	p
Screaming[a]	No	58	6	0.92[2]	1
	Yes	21	2		
Immobility[a]	0	11	0	- 0.85[1]	0.39
	90	21	2		
	120	14	2		
	180	33	4		
Alarm calling[a]	No	22	3	0.65[2]	0.68
	Yes	57	5		
Hissing[b]	No	52	4	2.14[2]	0.38
	Yes	18	3		
Attentiveness[b]	1	37	4	0.08[1]	0.94
	2	4	1		
	3	17	0		
	4	12	2		
Call inside nest box[b]	No	51	4	1.99[2]	0.40
	Yes	19	3		
Call 2 min after flight[b]	No	6	1	0.58[2]	0.52
	Yes	52	5		

[1]Mann-Whitney U test
[2]Fishers exact test
[a]Measured during the in-hand personality test
[b]Measured during the attentiveness test
Note: For description of the variables, see main text.

Appendix table 3.2.:

Table a. Comparing five continuous behavioural traits between male great tits that died or survived during the breeding season.

	Died	Survived		
Behavioural trait	**mean ± SD (N)**	**mean ± SD (N)**	**t/Z**	**p**
Breathing rate[a]	71.6 ± 17.5 (4)	66.0 ± 11.3 (77)	0.84[1]	0.41
Biting[a]	2.25 ± 2.63 (4)	2.82 ± 1.89 (77)	0.65[2]	0.52
Flight distance[a]	8.25 ± 1.71 (4)	8.53 ± 3.71 (77)	- 0.30[2]	0.77
Response to cage[b]	(1)	2.02 ± 2.16 (44)	-	-
Distance to predator[b]	(0)	2.28 ± 1.74 (56)	-	-

[1]Student's two sample t-test
[2]Mann-Whitney U test
[a]Measured during the in-hand personality test
[b]Too few individuals died for statistical testing.
Note: For description of the variables, see main text.

Table b. Comparing three categorical behavioural traits between male great tits that died or survived during the breeding season. Behavioural traits are measured in the in-hand personality test.

Behavioural trait		**Survived**	**Died**	**Z/OR**	**p**
Screaming	No	53	3	0.74[1]	1
	Yes	24	1		
Immobility	0	14	0	- 1.45[2]	0.15
	90	18	1		
	120	21	0		
	180	24	3		
Alarm calling	No	27	1	1.61[1]	1
	Yes	50	3		

[1]Fisher's exact test
[2]Mann-Whitney U test
Note: For description of the variables, see main text.

Appendix table 4:

Table a. Fitness variables in relation to age of female great tit parent.

Variable	First year mean ± SD (N)	Older mean ± SD (N)	Z/χ^2	p
Date of first egg (1 = 1 April)	28.55 ± 4.77 (44)	27.60 ± 5.62 (43)	1.15[1]	0.25
Clutch size	7.08 ± 1.29 (39)	7.72 ± 1.33 (43)	-2.38[1]	**0.017**
Incubation period (days)	16.97 ± 3.55 (37)	17 ± 3.66 (40)	-0.05[1]	0.96
Hatching date (1 = 1 May)	21.79 ± 5.13 (38)	21.12 ± 5.97 (43)	0.43[1]	0.67
No. nestlings hatched	7.32 ± 2.16 (38)	7.19 ± 2.22 (42)	0.43[1]	0.66
No. fledglings	4.29 ± 2.52 (35)	5.00 ± 2.35 (39)	-1.29[1]	0.20
No. nestlings dying	2.91 ± 3.02 (35)	2.32 ± 2.79 (38)	0.81[1]	0.42
% mortality of nestlings	36.23 ± 35.18 (35)	28.21 ± 28.95 (38)	0.81[1]	0.42
% females disappeared[a]	15.9 % (44)	14.3 % (49)	0.26[2]	0.61

[1]Mann-Whitney U test
[2]Pearson's chi-squared test
[a]No. females that disappeared in percentage of total number of females
Note: For description of the variables, see main text.

Table b. Fitness variables in relation to age of male great tit parent.

Variable	First year mean ± SD (N)	Older mean ± SD (N)	Z/OR	p
Date of first egg (1 = 1 April)	29.32 ± 5.16 (34)	26.91 ± 4.23 (46)	2.02[1]	**0.043**
Clutch size	7.13 ± 1.43 (31)	7.56 ± 1.27 (45)	-1.22[1]	0.22
Duration of incubation period	17.17 ± 2.98 (29)	17.00 ± 3.96 (42)	0.66[1]	0.51
Hatching date (1 = 1 May)	22.9 ± 5.05 (29)	20.13 ± 5.49 (45)	1.98[1]	**0.047**
No. nestlings hatched	7.03 ± 2.72 (29)	7.18 ± 1.96 (44)	-0.28[1]	0.78
No. fledglings	3.69 ± 2.51 (26)	5.00 ± 2.33 (42)	-2.28[1]	**0.022**
No. nestlings dying	3.27 ± 3.21 (26)	2.15 ± 2.63 (41)	1.51[1]	0.13
% mortality of nestlings	41 ± 35.15 (26)	27 ± 30.18 (41)	1.69[1]	0.091
% male disappeared[a]	11.8 % (34)	2.0 % (51)	0.15[2]	0.15

[1]Mann-Whitney U test
[2]Fisher's exact test
[a]No. males that disappeared in percentage of total number of males.
Note: For description of the variables, see main text.

Appendix table 5:

A summary of ANOVA/ANCOVAs where possible confounding variables are accounted for. The chosen variables for testing are based on relationships with $p < 0.05$ in a great tit population.

Dependent variable	Sex	Independent variable	F	p
Date of first egg[1]	Male	Distance to predator	4.21	**0.045**
		Age	4.32	**0.043**
Clutch size[2]	Female	Calling inside the box[a]	3.97	**0.050**
		Age	7.9	**0.006**
	Female	Attentiveness[a]	4.59	**0.005**
		Age	11.5	**0.001**

[1]ANCOVA.
[2]ANOVA.
[a]Measured during the attentiveness test
Note: For description of the variables, see main text.

Appendix table 6:

Spearman rank correlations between the female and male great tits' date of first egg and other fitness variables.

	Date of first egg (N=69)			
	Females		Males	
Variables	**r^s**	**p**	**r^s**	**p**
Clutch size	-0.19	0.11	-0.13	0.31
Incubation period	-0.02	0.90	-0.02	0.87
Hatching date	0.71	**< 0.001**	0.68	**< 0.001**
No. nestlings	0.02	0.89	-0.01	0.91
No. fledglings	-0.02	0.86	-0.06	0.62
No. nestlings dying	0.06	0.65	0.07	0.59
% mortality of nestlings	0.06	0.64	0.09	0.45
Female disappeared	0.09	0.44	0.04	0.77

Note: For description of the variables, see main text.